数値計算 期末試験問題

対象：計算機科学コース・ソフトウェア科学コース・数理科学コース

担当教員：大木 健太郎

2014 年 7 月 24 日 (木曜日 5 時限)

- 電卓および電卓機能を有する携帯電話の持ち込み可（PC やタブレット端末等は不可）
- 教科書・ノート等，およびそれに類する電子ファイルの閲覧は不可.
- 電卓機能以外の携帯端末等の利用は不可（とくにネットワーク機能は不可）.
- 試験中であっても教員が許可しないものは使用不可とし，試験終了時刻まで没収する.
- どうしても解けない場合，問題の意味を損なわない程度に簡単にして解いてよい．簡単にする旨を明記して解くこと．この場合，配点の 8 割を上限とし，作成した問題の程度に応じて減点する.

1. 区間 $[-1,1]$ において，次の関数を最大にする $x=x^* \in [-1,1]$ を求めたい.

$$S(x) = 3\left(x+\frac{1}{2}\right)e^{-x}$$

この解は $x^*=0.5$ で与えられる．これを多項式関数で近似することで数値的に求めたい．次の問に答えよ.

(a) $f(x)=\frac{dS(x)}{dx}$ とする．関数 f を，次の多項式

$$p_1(x) = a_1x+b_1$$

で最小二乗の意味で近似し，近似解 $-b_1/a_1$ を求めたい．これは,

$$\int_{-1}^{1}(f(x)-p(x))^2 = (A_1\theta_1-B_1)^\top A_1^{-1}(A_1\theta_1-B_1)+\int_{-1}^{1}f(x)^2dx - B_1^\top A_1^{-1}B_1,$$

$$\theta_1=\begin{bmatrix}a_1\\b_1\end{bmatrix},\ A_1=\begin{bmatrix}\int_{-1}^1 x^2dx & \int_{-1}^1 xdx\\ \int_{-1}^1 xdx & 2\end{bmatrix}=\begin{bmatrix}\frac{2}{3} & 0\\ 0 & 2\end{bmatrix},\ B_1=\begin{bmatrix}\int_{-1}^1 xf(x)dx\\ \int_{-1}^1 f(x)dx\end{bmatrix}=\begin{bmatrix}-3e+12e^{-1}\\ \frac{3}{2}e+\frac{9}{2}e^{-1}\end{bmatrix}$$

と変数 (a_1,b_1) について平方完成することができ，$A_1\theta_1=B_1$ を解けば求まる. $e\simeq 2.72$ とし，10 進 3 桁で $p_1(x)=0$ の解を求めよ.

(b) 多項式の次数を高めると，解の近似精度が改善される．これは Weierstrass の定理で保証される．多項式の次数を 2 として

$$p_2(x) = a_2x^2+b_2x+c_2$$

とすると，f の p_2 による最小二乗近似は，次の連立一次方程式の解である.

$$A_2\theta_2=B_2,\ \theta_2=\begin{bmatrix}a_2\\b_2\\c_2\end{bmatrix},\ A_2=\begin{bmatrix}\frac{2}{5} & 0 & \frac{2}{3}\\ 0 & \frac{2}{3} & 0\\ \frac{2}{3} & 0 & 2\end{bmatrix},\ B_2=\begin{bmatrix}-\frac{9}{2}e+\frac{81}{2}e^{-1}\\ -3e+12e^{-1}\\ \frac{3}{2}e+\frac{9}{2}e^{-1}\end{bmatrix}$$

Gauss の消去法を用いて θ_2 を求めよ（10 進 3 桁の浮動小数点数で表すこと）.

(c) 二次方程式 $ax^2+bx+c=0$ の根を求める計算アルゴリズムを1つ示し，上で得られた θ_2 から $f(x)=0$ の近似解を10進3桁の浮動小数点数で求めよ（反復法で行う場合，$x^{(0)}=0.6$ より始めるものとし，1ステップ後の解を示すこと）．また，2次の多項式による近似解の精度と Weierstrass の定理の主張について，論ぜよ（$x \geq x^*$ で $f(x)$ の傾きが小さくなることに注意せよ）．

2. 次のトピックの中から1つ選び，説明せよ．

(a) Gauss–Seidel 法について，そのアルゴリズムおよび収束が保証される行列の条件を述べよ．

(b) 連立一次方程式の共役勾配法のアルゴリズムと，その性質について述べよ．

(c) 非線形方程式の反復法における，縮小写像の原理について述べよ．

(d) 対称行列の3重対角化アルゴリズムである，Householder 法について述べよ．

(e) 代数方程式の根を求めるアルゴリズムである， DKA 法について述べよ．

3. $y=x^2$ と $x^2+xy=2$ の交点の x 座標を Newton 法を用いて求めたい．以下の問に答えよ．

(a) 初期値 $x^{(0)}=2$ で2ステップまで求めよ．計算の途中過程は10進3桁の浮動小数点数表示で行い，解も10進3桁の浮動小数点数で表せ.

(b) 浮動小数点数近似による誤差を考える．数値誤差のない場合との誤差が，2ステップでどの程度大きくなるか見積もりたい．方針を示せ.

4. Jacobi 法による実対称行列 $A=(a_{ij})\in\mathbb{R}^{n\times n}$ の対角化を考える．

(a) A の (p,q) 成分 a_{pq}，（$p\neq q$）を $PAP^\top$ なる変換で消去する直交行列 P は，

$$P_{ij}=\begin{cases}\cos\theta, & i=j=p,\ i=j=q\\ \sin\theta, & i=p,\ j=q\\ -\sin\theta, & i=q,\ j=p\\ 1, & i=j\neq p,\ i=j\neq q\\ 0, & \text{otherwise}\end{cases}$$

の形で与えられる．$\cos\theta$ と $\sin\theta$ を求めよ．

(b) Jacobi 法によって対角行列に収束することを示せ.

5. 連立一次方程式

$$Ax=b,\quad A\in\mathbb{R}^{n\times n},\ x,b\in\mathbb{R}^n$$

を解くには，A が正則行列であればよく，このとき厳密解 $x=A^{-1}b$ が存在する．しかし，数値計算では丸め誤差などの影響で，解が厳密解と大きく異なってしまう場合がある．次の2つの行列 A_1,A_2 の条件数 $\kappa(A):=\|A\|_2\|A^{-1}\|_2$ を求め，連立一次方程式の解きやすさについて論ぜよ．ただし，行列 A に対するノルム $\|A\|_2$ は，Euclid ノルムの誘導ノルムであるとする．また，どちらが解きにくいかを述べよ．

$$A_1=\begin{bmatrix}100 & 0\\ 0 & 0.01\end{bmatrix},\quad A_2=\begin{bmatrix}3 & 4\\ 4 & 3\end{bmatrix}$$

配点と採点時の注意

- 配点：
  - 1 (a) 5 点，　1 (b) 10 点，　1 (c) 10 点，
  - 2　　15 点，
  - 3 (a) 10 点，　3 (b) 10 点，
  - 4 (a) 10 点，　4 (b) 10 点，
  - 5　　20 点
- 計算過程から明らかに計算ミス，あるいは書き間違いと分かるは，減点しない．
- 浮動小数点数の表記は，$.198 \times 10^2$ でも 19.8 でもどちらでも可とした（講義でも表記の簡略化のために後者を採用したので）．
- 特に指示していない問題で計算途中で 10 進 3 桁の浮動小数点を用いた場合，計算結果の精度が許容範囲内（丸め誤差の伝播の上限値以内）ならば正解とした．

解答例

1. (a) 次の方程式

$$a_1 = \frac{3}{2}\left(-3e + 12e^{-1}\right) \simeq -5.6104,\ b_1 = \frac{1}{2}\left(\frac{3}{2}e + \frac{9}{2}e^{-1}\right) \simeq 2.8664.$$

よって，解は，$x^* \simeq 0.5109 \simeq 0.511$.

(b) 解くべき代数方程式は $A_2\theta_2 = B_2$ であり，３行目の第１列を消去すれば，後退代入より全ての解が求められる．

$$\theta_2 \simeq \begin{bmatrix} 4.25, & -5.61, & 1.45 \end{bmatrix}^\top$$

(c) 例えば Newton 法を用いればよい．$f(x) = ax^2 + bx + c$ とし，その微分 f' も用いると，Newton 法は次のアルゴリズムで表される．

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, \quad k = 0, 1, 2, \cdots.$$

１ステップのみ計算すると，

$$\begin{aligned} x^{(1)} &= x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})} \\ &= 0.6 - \frac{0.36a + 0.6b + c}{1.2a + b} \simeq -0.160 \end{aligned}$$

である．(このアルゴリズムを続けると，約 0.35 付近に収束する)

Weierstrass の定理は，多項式の次数を大きくすれば，任意の連続関数の近似が任意の精度でできることを述べている．ここで任意の精度とは，関数と多項式の差に対し，講義では閉区間の二乗積分の意味で，教科書では max ノルムの意味で与えている．どちらの場合も，零点の近さと精度とは直接関係がなく，とくに変化が小さい関数に対しては零点が大きく異なる場合もありうる．この場合，$f(x)$ の傾きは $x \geq x^*$ で小さいため，f の零点の位置と近似多項式の零点の位置は大きく異なる可能性がある．問題の関数の零点付近の多項式近似の様子を図 1 に示す．

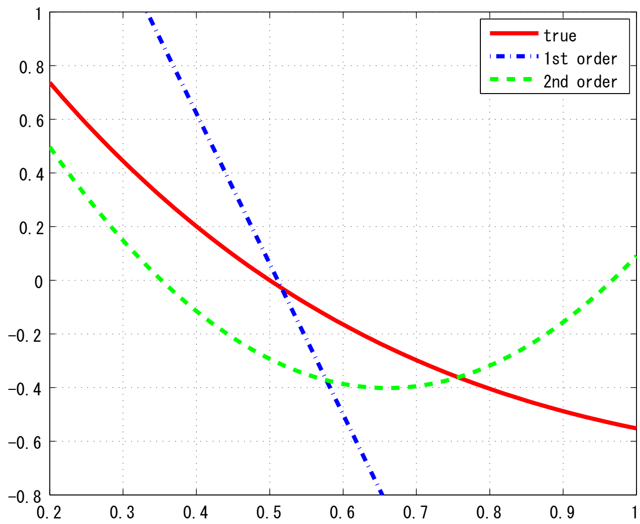

図 1: 関数 f とその多項式近似.

2. (a) Gauss–Seidel 法とは，連立一次方程式 $Ax = b$ を次の形で回帰的に解く数値計算法の 1 つである．

$$x^{(k+1)} = Mx^{(k)} + Nb \tag{1}$$

ただし，$\det(A) \neq 0$ とする．ここで M と N は，A をそれぞれ対角行列 D，狭義下三角行列 E，狭義上三角行列 F で分解した

$$A = D + E + F$$

を用いて

$$M = -(D+E)^{-1}F, \quad N_G = (D+E)^{-1}$$

で定義される．丸め誤差のない場合，このアルゴリズムが収束するための必要十分条件は，M の固有値の絶対値が 1 未満になることである．例えば A が正定値行列の場合，数値誤差がなければ真の解へ収束する．

(b) 共役勾配法とは，正定値行列 A を係数とする連立一次方程式 $Ax = b$ の解を適当な初期値を用いて

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

なる差分方程式の解として構成する反復法の一種である．ここで $p^{(k)} \in \mathbb{R}^n$ は修正方向，$\gamma_k \in \mathbb{R}$ は修正の大きさを表し，それぞれ

$$p^{(k)} = r^{(k)} + \beta_{k-1}p^{(k-1)},\ r^{(k)} := (b - Ax^{(k)}),\ \beta_{k-1} := -\frac{(p^{(k-1)})^\top Ar^{(k)}}{(p^{(k-1)})^\top Ap^{(k-1)}},\ \alpha_k = \frac{(r^{(k)})^\top r^{(k)}}{(p^{(k)})^\top Ap^{(k)}}$$

で定められる．これを $k = 0$ から繰り返せばよい．共役勾配法は直交性に基づいているため，丸め誤差に強い．また，丸め誤差がない場合，n 階のステップで厳密解に収束するという性質を持っている．

(c) ノルム $\|\bullet\|$ が定められた線形空間 X の閉領域 R が与えられているとする．任意の $x \in R$ に対して $\phi(x) \in R$ を満たす写像 ϕ が，ある $q \in [0,1)$ に対して

$$\|\phi(x) - \phi(y)\| \leq q\|x - y\|, \quad \forall x, y \in R$$

を満たすとき，ϕ を R 上の縮小写像という．縮小写像の原理とは，反復法

$$x^{(k+1)} = \phi(x^{(k)}), \quad k = 0, 1, 2, \cdots$$

の ϕ が閉領域 R で縮小写像であり，初期値が R 内にあるとき，反復法による解は収束し，収束先は R 内に一意に存在する．

(d) Householder 法とは，Givens 変換と呼ばれる直交変換を用いて，対称行列を３重対角化する手法である．主に中規模〜大規模な行列の固有値問題の前処理に用いられる．直交行列 P_i を用いて

$$A^{(1)} = P_1^\top A P_1 = \begin{bmatrix} * & * & 0 & \cdots & 0 \\ * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \cdots & * \end{bmatrix}$$

$$A^{(2)} = P_2^\top A^{(1)} P_2 = \begin{bmatrix} * & * & 0 & \cdots & & 0 \\ * & * & * & 0 & \cdots & 0 \\ 0 & * & * & \cdots & & * \\ 0 & 0 & * & \cdots & & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & * & \cdots & * & \end{bmatrix}$$

と左上から順次，３重対角化する．ここで P_i の選び方は，$A = (a_{ij})$ に対し

$$x = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix},\ y = \begin{bmatrix} a_{11} \\ \sqrt{\sum_{j=2}^n a_{j1}^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ u_1 = \frac{x-y}{\|x-y\|_2},\ P_1 = I - 2u_1u_1^\top$$

とすればよい．次のステップでは，

$$x = \begin{bmatrix} a_{12}^{(1)} \\ a_{22}^{(1)} \\ a_{32}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{bmatrix},\ y = \begin{bmatrix} a_{12}^{(1)} \\ a_{22}^{(1)} \\ \sqrt{\sum_{j=3}^n (a_{j2}^{(1)})^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ u_2 = \frac{x-y}{\|x-y\|_2},\ P_2 = I - 2u_2u_2^\top$$

とすればよい．これを繰り返すと，$n-1$ ステップで三重対角化が可能である．

(e) DKA 法（Durand–Kerner–Aberth 法）とは，代数方程式の数値解法の１つで，Newton 法を改良したものである．n 次の多項式

$$p_n(z) = a_0z^n + a_1z^{n-1} + \cdots a_{n-1}z + a_n$$

について，

$$z_i^{(k+1)} = z_i^{(k)} - \frac{p_n(z_i^{(k)})}{a_0\prod_{j\neq i}(z_i^{(k)} - z_j^{(k)})},\quad i = 1, 2, \cdots, n$$

とし，初期値を

$$z_i^{(0)} = -\frac{a_1}{n} + r_0 \exp\left(\sqrt{-1}\left(\frac{2(i-1)\pi}{n} + \frac{\pi}{2n}\right)\right)$$

と選ぶアルゴリズムのことで，重根がなく数値誤差もない場合は真値に収束する．

3. (a) $y = x^2$ を $x^2 + xy = 2$ に代入すると，交点の座標は，

$$f(x) := x^2 + x^3 - 2 = 0$$

を満たす x である．Newton 法は

$$
\begin{aligned}
x^{(k+1)} =& x^{(k)} - \frac{f(x^k)}{f'(x^{(k)}} \\
=& x^{(k)} - \frac{(x^{(k)})^2 + (x^{(k)})^3 - 2}{2x^{(k)} + 3(x^{(k)})^2}
\end{aligned}
$$

である．初期値を $x^{(0)} = 2$ とすると，

$$
\begin{aligned}
x^{(1)} =& 2 - \frac{4+8-2}{4+12} \\
=& 2 - \frac{10}{16} = 2 - 0.625 \\
\simeq& 2 - 0.63 = 1.37
\end{aligned}
$$

を得る．$x^{(1)} = 1.37$ として，

$$
\begin{aligned}
x^{(2)} =& 1.37 - \frac{1.8769 + 2.5714 - 2}{2.74 + 5.6307} \\
\simeq& 1.37 - \frac{1.88 + 2.57 - 2}{2.74 + 5.63} = 1.37 - \frac{2.45}{8.37} \\
\simeq& 1.37 - 0.293 \simeq 1.37 - 0.29 = 1.08
\end{aligned}
$$

を得る．

(b) 反復法を

$$
x^{(k+1)} = \phi(x^{(k)})
$$

とおくと，数値誤差のある場合は

$$
y^{(k+1)} = \phi(y^{(k)}) + \xi^{(k)}
$$

と表せる．ここで $\xi^{(k)}$ は丸め誤差などによって生じる数値誤差を表す．考えている閉領域 R で ϕ が縮小写像であるとすると（そうでなければ，得られた結果が真値である保証がない），適当な $q \in [0, 1)$ が存在して，

$$
\|\phi(x) - \phi(y)\| \le q\|x - y\|, \quad \forall x, y \in R
$$

を満たす．このとき，

$$
\begin{aligned}
\|x^{(k+1)} - y^{(k+1)}\| =& \|\phi(x^{(k)}) - \phi(y^{(k)}) - \xi^{(k)}\| \\
\le& \|\phi(x^{(k)}) - \phi(y^{(k)})\| + \|\xi^{(k)}\| \\
\le& q\|x^{(k)} - y^{(k)}\| + \|\xi^{(k)}\|
\end{aligned}
$$

である．任意の k に対して $\|\xi^{(k)}\| \le \epsilon$ を満たす $\epsilon > 0$ が存在する場合，真値との誤差は最悪でも

$$
\frac{\epsilon}{1-q}
$$

になる．したがって，閉領域で縮小写像であることと，q を評価することが，誤差の評価につながる．詳細には各ステップ毎に計算することもできる．

4. (a) n 次直交行列の集合を $\mathcal{O}(n)$ とする．Jacobi 法とは，問 4 のように適切に添え字 p, q を選び，

$$
\begin{aligned}
& P_i \in \mathcal{O}(n), \\
& P_1^\top A P_1 \mapsto P_2^\top P_1^\top A P_1 P_2 \mapsto \cdots \mapsto P_k^\top \cdots P_1^\top A P_1 \cdots P_k \\
& \to \text{対角行列}
\end{aligned}
$$

とするものである．A の (i,j) 成分 a_{ij} とし，

$$B = (b_{ij}) := P^\top AP$$

とおく．ここで B の (p,q) 成分は

$$\begin{aligned}
b_{pp} =& \frac{a_{pp}+a_{qq}}{2} - \frac{a_{pp}-a_{qq}}{2}\cos(2\theta) - a_{pq}\sin(2\theta), \\
b_{qq} =& \frac{a_{pp}-a_{qq}}{2} - \frac{a_{pp}-a_{qq}}{2}\cos(2\theta) + a_{pq}\sin(2\theta), \\
b_{pq} =& b_{qp} = \frac{a_{pp}-a_{qq}}{2}\sin(2\theta) + a_{pq}\cos(2\theta)
\end{aligned}$$

であり，Jacobi 法は

$$b_{pq} = 0$$

となるように θ を決めるアルゴリズムである．θ は

$$\tan(2\theta) = \frac{-2a_{pq}}{a_{pp}-a_{qq}}$$

を満たすように定めればよい．

$$\cos(2\phi) = \frac{|a_{pp}-a_{qq}|}{\sqrt{4a_{pq}^2 + (a_{pp}-a_{qq})^2}}$$

より，倍角の公式から

$$\begin{aligned}
\cos(\phi) =& \sqrt{\frac{1}{2}(1+\cos(2\phi))}, \\
\sin(\phi) =& \mathrm{sgn}\left(-a_{pq}(a_{pp}-a_{qq})\right)\sqrt{\frac{1}{2}(1-\cos(2\phi))}
\end{aligned}$$

を得る．

(b) 行列の非対角成分の二乗和を

$$F = \sum_{i\neq j} a_{ij}^2, \quad \tilde{F} = \sum_{i\neq j}(P^\top AP)_{ij}^2 = \sum_{i\neq j} b_{ij}^2$$

とする．ここで先の問題で示した以外の B の要素は，

$$\begin{aligned}
b_{ij} =& a_{ij},\ (i,j \neq q,p), \\
b_{pj} =& b_{jp} = a_{pj}\cos(\theta) - a_{pj}\sin(\theta),\ (j\neq q,p) \\
b_{qj} =& b_{jq} = a_{qj}\cos(\theta) + a_{qj}\sin(\theta),\ (j\neq q,p)
\end{aligned}$$

である．

$$\begin{bmatrix} b_{pj} \\ b_{qj} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} a_{pj} \\ a_{qj} \end{bmatrix}$$

より，

$$\begin{aligned}
b_{pj}^2 + b_{qj}^2 = \begin{bmatrix} b_{pj} & b_{qj} \end{bmatrix}\begin{bmatrix} b_{pj} \\ b_{qj} \end{bmatrix} &= \begin{bmatrix} a_{pj} & a_{qj} \end{bmatrix}\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}^\top \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} a_{pj} \\ a_{qj} \end{bmatrix} \\
&= \begin{bmatrix} a_{pj} & a_{qj} \end{bmatrix}\begin{bmatrix} a_{pj} \\ a_{qj} \end{bmatrix} = a_{pj}^2 + a_{qj}^2
\end{aligned}$$

である．したがって，

$$\tilde{F} - F = -2a_{pq}^2$$

を得る．したがって，$a_{pq} \neq 0$ となる添え字を選び続ければ，対角行列に収束する．

5. A_1 と A_2 は対称行列なので，条件数は固有値の絶対値の最大値と最小値の逆数の積になる．A_1 と A_2 の条件数は，それぞれ

$$\kappa(A_1) = 10^4, \quad \kappa(A_2) = 7$$

である．一般論として，条件数の大きな行列は数値誤差によって計算結果が真値と大きく異なることがある．しかし，与えられた A_1 は対角行列であり，数値計算で実際に扱うのは

$$100x_1 = b_1, \quad 0.01x_2 = b_2$$

なる 1×1 の 2 つの連立一次方程式であり，それぞれ条件数 1 である．とくに 10 進数では指数部を変化させるだけなので，A_1 の方が解きやすい．